

基于 OPENCV 的手势识别系统的设计与实现

陆海虹

(西安邮电大学 物联网与两化融合研究院, 西安 710072)

摘要: 人与计算机的交互技术是一种新型的计算机技术, 且逐渐演变为一种主流技术和计算机领域的技术热点; 为了能够更好地识别手势和跟踪手势的运动轨迹, 提出了基于 OPENCV 的手势识别系统, 系统引入了 OPENCV 计算机视觉库, OPENCV 作为优秀的计算机视觉库, 为设计的实现提供了便捷的代码, 利用 OPENCV 技术中的图像处理算法, 首先通过摄像头采集数据图像, 并对采集到的图像进行一系列的缩放, 去噪以及锐化等处理, 然后对人体手势建立肤色模型, 然后经过灰度阈值化来转换成二值图像, 得到手轮廓的数据图像后, 采用轮廓匹配方法识别出手型; 最后通过 10 种基本的手势模型对比验证了本系统具有一定的实时性, 并且识别率可以达到 95% 以上。

关键词: 手势识别; 人机交互; OPENCV; 轮廓匹配; 肤色模型

Design and Implementation of Gesture Recognition System Based on OPENCV

Lu Haihong

(Institute of IoT & IT-based Industrialization, Xi'an University of Post and Communication, Xi'an 710072, China)

Abstract: Interaction of human and computer is a computer model, and gradually evolved into a hot technology in a mainstream technology and computer science. In order to trajectory identification and tracking of gestures can be better, the gesture recognition system based on OPENCV is proposed, the system introduces the OPENCV computer vision library, computer vision library OPENCV as excellent, provides convenient code for the realization of the design, the use of image processing algorithms in OPENCV technology, is the first to establish the skin color model of human gesture then, through gray level thresholding to convert binary value image, then the morphological transformation, finally obtains the data image hand contour, the contour matching method, the final recognition hand type, and can effectively analyze the motion trajectory. The experiments verify the effectiveness and practicability of the system.

Keywords: hand gesture recognition; human-computer interaction; OPENCV; color model; contour matching

0 引言

人机的交互活动最早诞生是为了能更好更方便的控制计算机, 输入信息。最常用的也是最突出的就是利用键盘和鼠标进行人与计算机的信息交换, 但是这也成为了人机交互发展的瓶颈。当前广泛使用的 WIMP 图形用户接口, 因为从用户到计算机的通信是串口的, 用户不能以手势、语音进行人机交互。一般计算机处理过程将抛弃传统的交互设备, 诸如键盘、鼠标等传统设备, 近而需要一些自动化程度较高的设备来采集记录, 这里面包括各种感知器和采集器(摄像机等)^[1]。

对于手势识别技术相关研究具有代表性的有, 1991 年富士通实验室完成了对 46 个手语符号的识别工作^[2]。Davis 和 Shah 将戴在指间具有高亮标记的视觉手套的手势作为系统的输入, 可以识别 7 种手势^[3-4]。目前最先进的发展是名为“真实手套”的技术, 目前已经可以检测手指的弯曲指数和空间定位指数, 使三维交互手段变得更加的真实自然^[5-6]。但是不可避免的是这种技术造价必然昂贵, 即使技术完善也并不能投入大规模的使用。所以基于计算机视觉的手势识别技术依旧是未来的发展趋势^[7-8]。

本文为了能够更好地识别和跟踪手势的运动轨迹, 提出了基于 OPENCV 的手势识别系统, 系统引入了 OPENCV 计算机视觉库, 采用轮廓匹配方法, 最终识别出手型, 并可以有效

分析出运动轨迹。最后实验验证了系统的有效性和实用性。

1 系统总体设计

1.1 系统总体结构

本体系统大体上分为 3 个部分。面向用户的参数调整模块, 试用 MFC 程序配合 OPENCV 库调用摄像头函数进行视频提取, 调整函数, 该部分主要的功能是为系统提供原始数据, 并且让用户进行参数调整以应对不同光线环境和其他影响下, 能够更好的提高识别率。第二部分是肤色提取主模块, 该部分的功能对应着系统需要分析时的轮廓提取步骤, 模板匹配只能针对手势的二值图像进行匹配, 我已我们要通过一定的算法对初始输入数据进行处理。包括预处理的降噪, 肤色提取。和最后的轮廓提取。最后一部分是手势识别主模块, 程序运行时, 当输入数据已经完成处理之后, 将处理后的数据与预想设定的模板数据进行匹配。系统结构图如图 1 所示。

1.2 数据采集与图像预处理模块

要进行手势信息的获取, 首先需要拥有一定的硬件, 如数据手套、摄像头等。对于软件方面则要安装相应的驱动程序和了解它们的 API。本文利用 OPENCV 具有简洁的外部接口的优点, MFC 类库和 OpenCV 可以将摄像头的源数据或数据的处理情况呈现到控件中。

由于摄像头的图像尺寸千变万化, 不一定能满足手的关键点位置的精准度或者图片太大造成延迟和卡顿, 而且可能包含了大量的阻碍识别的信息(如点状噪声等), 因而需要一些对这些采集到的图像进行一系列的预处理操作。

收稿日期: 2014-07-17; 修回日期: 2014-10-17。

作者简介: 陆海虹(1972-), 男, 上海人, 硕士, 讲师, 主要从事多媒体通信方向的研究。

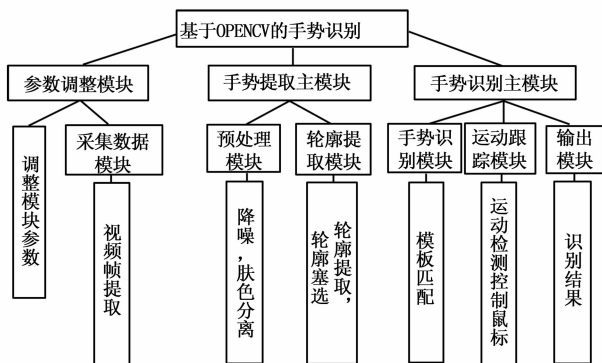


图 1 系统结构图

1.2.1 图像缩放

由于通过摄像头采集到的图像非常大，对系统的处理时间造成延时，所以必须要对其进行缩放操作，即需要缩放到一个合理的大小，如 320×240 ，为了加强精确度有条件的可以加到 640×480 。OpenCV 提供了一个函数 `cvResize` 可以实现这个功能。如原 `IplImage*` 图片是 `m_Captured`，目标图片是 `m_Frame`，则函数可写成：

```
cvResize(m_Captured, m_Frame(m_Captured->width) > (m_Frame->width)? CV_INTER_LINEAR:CV_INTER_AREA);
```

当原图片比目标图片小时应使用双线性或双三次插值的放大算法，来使得手的运动轨迹连续变化，反之，则采用平均值的缩小算法，来减少波纹和边缘的不连续现象。

1.2.2 图像降噪处理

图像噪声按产生的原因可分为外部噪声和内部噪声。外部噪声是指由于系统外部干扰以电磁波或由电源串进系统内部而引起的噪声。如电气设备，天体放电现象等引起的噪声，而这种噪声可能就是高斯噪声、脉冲噪声等多个噪声合成累计的。内部噪声主要是由光和电的基本性质所引起的噪声。如电流的产生是由空穴或电子的集合，定向运动所形成的，而这些粒子运动的带有随机性，产生了散粒噪声；而导体中也有自由电子，它们会进行无规则热运动，形成热噪声。为了消除这些随机因素，本文通过中值滤波或均值滤波的方法法进行滤波。对应的 OpenCV 函数是：

```
void cvSmooth( const CvArr* src, CvArr* dst, int smoothtype=CV_GAUSSIAN, int param1=3, int param2=0, double param3=0, double param4=0)。
```

其中 `smoothtype` 是 `CV_GAUSSIAN` 时将图像进行大小为 `param1 × param2` 的高斯卷积的均值滤波。而 `smoothtype` 是 `CV_MEDIAN` 时将图像进行大小为 `param1 × param1` 的中值滤波，注意为 `param1` 奇数，且不大于 7，其效果结果如图 2 所示。



(a)原图 (b)CV_GAUSSIAN (c)CU_MEDIAN

图 2 CvSmooth 的不同 smoothtype 的比较

1.2.3 图像锐化处理

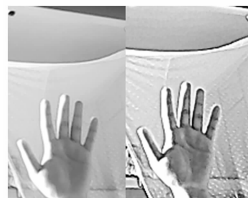
一般来说，经过图像的降噪处理后会显得比较朦胧，有时候需要进行 USM 锐化方可实现更好的边缘提取。为此可以先

用 `cvSmooth`，`smoothtype` 选 `CV_GAUSSIAN` 进行均值滤波到一个临时图像，然后按一定阈值将原图像与临时图像的像素反差扩大即可，反差扩大的公式为：

$$X = k(x - \bar{x}) + \bar{x}$$

式中， x 表示原图像的亮度， \bar{x} 表示临时图像的亮度， k 表示锐化的强度 ($k > 1$)， X 表示目标图像的亮度。

而通过原图像与临时图像的色差的阈值的调整，可以减少一些不需要锐化的区域，但如上所述，会造成边缘突兀和不连续，为此也可以将位于原图像与临时图像的亮度差在阈值内部部分的锐化强度，按这个亮度差线性减弱。如图 3 所示。



(a)锐化前 (b)锐化后
图 3 锐化前后的图像比较

2 手势区域特征提取

基于计算机视觉的手势识别的首要任务就是要将手和背景区分开，而背景环境往往会十分复杂，前面通过一系列的降噪和锐化实现了冗余信息的剔除和重要信息的强化，从而为接下来的操作作准备，下面主要给出如何尽可能准确地将手所在的区域从复杂的背景中提取出来。

肤色模型建立：

人眼对于色彩的敏感度要强于其他参数，因而在复杂环境中人眼接收彩色图像较黑白图像更为快速，信息量更大。虽然肤色人人均有不同，但是整体手势中，肤色占去的比例十分巨大，尤其是排除了光线变化影响之后，这种效果更加明显，所以采用 HSV 空间进行手势分割效果更好，更加简便，可应用于实时手势提取。

HSV 是简化了的孟塞尔色彩空间，以色调 (H)，饱和度 (S)，亮度 (V) 为 3 种参数来表示色彩。

从 RGB 色彩空间映射到 HSV 色彩空间的方法如下：

$$s = \begin{cases} \frac{MAX(R,G,B) - MIN(R,G,B)}{MAX} & \\ 0 & \end{cases} \quad (1)$$

当 $MAX(R,G,B) \neq 0$

$$H = \begin{cases} \frac{(-B + G) * \pi / 3}{MAX(R,G,B) - MIN(R,G,B)} & \text{当 } R = MAX(R,G,B) \\ \frac{(-R + G) * \pi / 3}{MAX(R,G,B) - MIN(R,G,B)} & \text{当 } G = MAX(R,G,B) \\ \frac{(-G + R) * \pi / 3}{MAX(R,G,B) - MIN(R,G,B)} & \text{当 } B = MAX(R,G,B) \end{cases} \quad (2)$$

$$V = MAX(R,G,B) \quad (3)$$

由于人手的主要肤色集中在红色区域，表现在图像中就是 $R > G > B$ 。据此原理将其简化后得到以下要素表达式：

$$S = \frac{R - B}{R}$$

$$H = \frac{(-B + G) * \pi / 3}{R - B}$$

$$V = R \quad (4)$$

在 OPENCV 中具体实现如下：

设 m_Frame 为 IplImage * 的图片，src 为它的 imageData 的开头，则鼠标在图像的坐标 (x0, y0) 上的像素指针为 mByte=src+x0 * (m_Frame->nChannels) +y0 * (m_Frame->widthStep)；

关于这一像素的 COLORREF 颜色：

colour= (int) mByte [0] <<16 | (int) mByte [1] <<8 | (int) mByte [2]；

需要注意的是，图像不一定充满整个控件，如果在空间外将会出现一些越界的情况，为此需要判断这个点是否在图像上：

if (x0>=0 && x0<m_Frame->width | | y0>=0 && y0<m_Frame->height)

OpenCV 中实现上述方法的函数为 void cvCvtColor (const CvArr * src, CvArr * dst, int code)，其中 src 输入的是原图像指针，而 IplImage 继承了 CvArr，有时需要将用整数信息存储的图片转换为一定精度的浮点图；dst 输入的是目标图像的指针；code 是色彩空间转换的模式，该 code 来实现不同类型的颜色空间转换。比如当 code 选用 CV_BGR2_HSV 时，对于 8 位图，需要将 RGB 值归一化到 0-1 之间。这样得到 HSV 图中的 H 取值范围才是 [0, 360)，S 和 V 的取值范围是 [0, 1]。

3 系统设计与实现

3.1 系统环境设置

本文设计的手势识别系统 GestureTest 的开发环境采用 Visual C++ 2010，系统类型采用了 MFC 基于对话框的应用程序，第三方库选用 OpenCV 作为识别的辅助工具。所以必须先进行工程环境的配置：

1) 安装 OpenCV2.0，并在其安装目录下的 include 和 lib 下的所有文件复制到 C:\Program Files\Microsoft Visual Studio 10.0\VC\下的同名目录中。

2) 打开所建立的工程，选择项目-GestureTest 属性，配置选“所有配置”，并在其下的配置属性-链接器-输入中输入 cxcv200d.lib; cv200d.lib; highgui200d.lib;% (AdditionalDependencies)

3) 将 OpenCV 安装目录下的 bin 目录里的 cv200d.dll、cxcv200.dll、cxcv200d.dll 和 highgui200d.dll 复制到项目生成的可执行文件的相同目录下。如果不能运行，则复制该 bin 目录里文件名不带 d 结尾（不含扩展名）的相应文件，并将程序旁 dll 文件的文件名后加上 d 即可。

建立的项目主要包含以下文件：

- GestureTest.h
- GestureTestDlg.h
- source.h
- stdafx.h
- targetver.h
- GestureTest.cpp
- GestureTestDlg.cpp
- stdafx.cpp

其中的 GestureTestDlg.cpp，即 CGestureTestApp 类在 InitInstance () 方法的 dlg.DoModel 后需要加上

```
if(dlg.m_Video! =NULL){
cvReleaseCapture(&(dlg.m_Video));
```

} 用来在程序结束后释放视频指针变量 m_Video，否则程序在关闭后无法彻底退出。

主要的实现方法将在 CGestureTestDlg 类中实现。在构造函数中将成员变量和控件变量赋上初始值，并在 OnInitDialog 中进行如图 4 所示的操作。

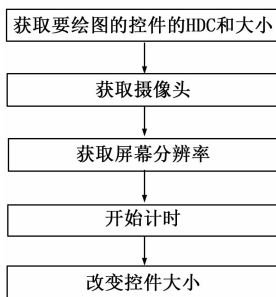


图 4 程序的初始化流程

3.2 手势识别实验

本系统定义了 10 种基本的手势模型，将该 10 个手势模板作为匹配模板进行手势识别，如图 5 所示。

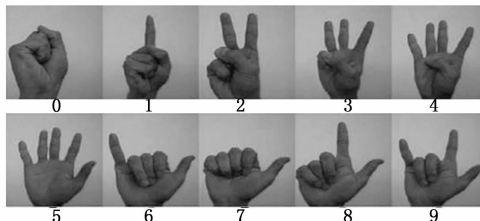


图 5 10 中手势模型

对应的该 10 种手势模型经过一系列图像分割锐化等处理后如图 6 所示。

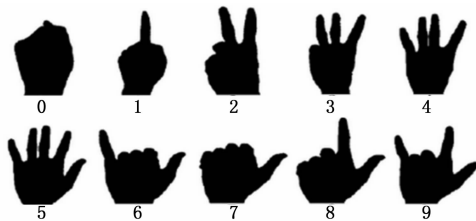


图 6 10 种图像处理后

本文在具体实验中，邀请 10 个实验者在数字视频摄像头前面做出相应的手势图像，通过摄像头采集这些手势，并把这些数据保存下来作为测试样本，每种手势做 8 次，共计 80 副，每种手势取 5 副，共取 50 副作为测试集进行测试。也就是找到特征手势与模版手势测试的距离最小的手势，作为最优的识别率。测试结果如表 1 所示，这里由于篇幅限制，本文也就给出一种摄像头采集的手势图像，如图 7 所示。



图 7 摄像头采集的手势与识别的手势

表 1 识别率

手势	1	2	3	4	5	6	7	8	9	10
识别率/(%)	99	100	95	97	99	98	92	96	100	99

本文将提出的方法应用在了无线通信发射控制系统上，首先通过用户编程来获取数据，然后经过 Zig2Serial USB 转串口信号转化板传给无线通讯模块 Zig-100 发射手势控制方向编码，接下来小车上的 Zig-100 无线通讯模块接收手势控制方向编码，传给小车控制系统，小车根据手势编码即可进行运动。

4 结束语

当今的人机交互技术已经成为了制约计算机应用和发展的瓶颈之一。本文利用 OPENCV 更好的图像处理技术以及实时的交互界面，设计了一种基于 OPENCV 的手势识别系统，并在系统中给出了 10 种简单手势的识别率，这说明本系统具有一定的实用性。可以推广到一些具体的应用领域。

参考文献:

[1] 关 然, 徐向民, 罗雅楠, 等. 基于计算机视觉的手势检测识别技术 [J]. 计算机应用与软件, 2013, 30 (1): 155-159.

[2] 陈锻生, 刘政凯. 肤色检测技术综述 [J]. 计算机学报, 2006, 29 (2): 194-207.

[3] 黄季冬. 动态手势识别技术研究及实现 [D]. 武汉: 华中科技大学, 2012.

[4] 万海根, 肖海英, 邹 松, 等. 面向新一代大众游戏的手势交互技术 [J]. 计算机辅助设计与图形学学报, 2011, 23 (7): 1159-1165.

[5] Pavlovic V, Sharma R, Huang T. Visual interpretation of hand gestures for human-computer interaction: A review [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 19 (7), 677-695.

[6] Quek, F. Gesture, speech, and gaze cues for discourse segmentation [A]. In IEEE Conference on Computer Vision and Pattern Recognition [C]. 2000. 247-254.

[7] 任雅祥. 基于手势识别的人机交互发展研究 [J]. 计算机工程与设计, 2006, 27 (7): 1201-1204.

[8] 李 健, 陈长明, 廖秋筠. 一种新的基于 OpenCV 的立体视觉三维测量方法 [J]. 计算机测量与控制, 2009 (12): 2403-2406.

(上接第 1645 页)

表 2 逆模型辨识适应度值变化

迭代数	r_1	r_2	r_3	r_4	r_5	α	β	$Fitness_{best}$
300	1.5e+03	109.165	2.6546	0.0103	0.4188	4.2e-04	4.0e-04	1.1288e+07
600	6.8e+03	85.326	1.1656	0.0012	0.1567	4.4e-04	4.2e-04	3.5625e+04
900	5.2e+04	72.432	0.8469	0.0089	0.0849	5.1e-04	4.6e-04	167.6033
1200	5.6e+04	73.062	0.6613	0.0125	0.0518	8.2e-04	5.9e-04	6.1856

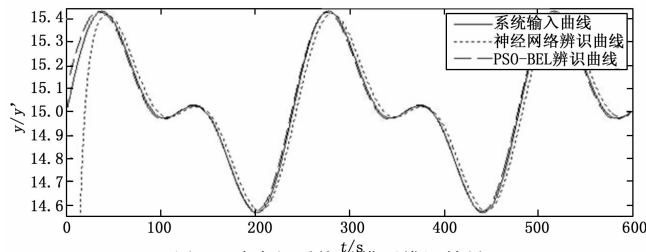


图 7 球磨机系统逆模型辨识结果

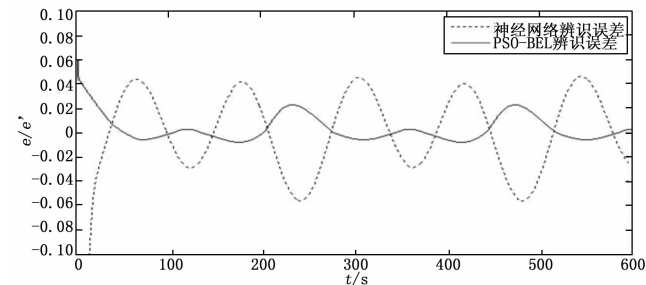


图 8 球磨机系统逆模型辨识误差

5 结论

本文通过粒子群优化算法优化最佳感官输入函数、情感暗示函数等参数，使大脑情感学习模型 BEL 在球磨机系统辨识中达到最佳效果。实验结果表明本文方法可使系统辨识精度提高，降低误差。对于实现球磨机基于模型控制方法的实现具有重要的意义。

参考文献:

[1] 李 丹. 球磨机制粉系统建模及其解耦控制研究 [D]. 吉林: 东北电力大学, 2014.

[2] 黄 希, 王 恒, 马海波, 等. 融合筒体振动的球磨机系统瞬态响应建模研究 [J]. 矿山机械, 2013, 41 (7): 63-67.

[3] Decety J, Michalska K J, Kinzler K D. The contribution of emotion and cognition to moral sensitivity: a neurodevelopmental study [J]. Cerebral Cortex, 2012, 22 (1): 209-220.

[4] Jain A, Tayal D, Sehgal N. Brain Emotional Learning Based Intelligent Controller and its Application to Jacketed Continuous Stirred Tank Reactor [J]. International Journal, 2013.

[5] Kennedy J. Particle swarm optimization [M]. Encyclopedia of Machine Learning. Springer US, 2010; 760-766.

[6] 郑太雄, 李 锐. 基于粒子群算法的汽车 ABS 控制器参数的优化设计 [J]. 机械科学与技术, 2007, 26 (9): 1172-1174.

[7] 张耀军, 谌昌强. 改进量子 PSO 算法的可约束车辆路径优化 [J]. 计算机测量与控制, 2014, 22 (9): 2875-2878.

[8] 徐玉杰. 粒子群算法的改进及应用 [D]. 南京: 南京师范大学, 2013.

[9] 叶志伟, 周 欣, 郑肇葆, 等. 产生“Tuned”模板的混沌粒子群算法 [J]. 武汉大学学报: 信息科学版, 2013 (1): 10-14.

[10] 曾 毅, 朱旭生, 廖国勇. 一种基于邻域空间的混合粒子群优化算法 [J]. 华东交通大学学报, 2013, 30 (3): 44-49.

[11] 李国柱. 基于 PSO 的测距模型参数估计三维定位算法 [J]. 计算机测量与控制, 2014, 22 (9): 2853-2855.

[12] 董毅峰, 王雪瑶. 基于神经网络的闭路球磨机产量控制方法研究 [J]. 煤矿机械, 2007, 28 (5): 40-42.